Human Movement Datasets: An Interdisciplinary Scoping Review

TEMITAYO OLUGBADE, University College London, United Kingdom
MARTA BIEŃKIEWICZ, EuroMov Digital Health in Motion, Univ. Montpellier IMT Mines Ales, France
GIULIA BARBARESCHI, University College London, United Kingdom
VINCENZO D'AMATO, LUCA ONETO, and ANTONIO CAMURRI, Università di Genova, Italy
CATHERINE HOLLOWAY, University College London, United Kingdom
MÅRTEN BJÖRKMAN, KTH Royal Institute of Technology, Sweden
PETER KELLER, Western Sydney University, Australia
MARTIN CLAYTON, Durham University, United Kingdom
AMANDA C DE C WILLIAMS and NICOLAS GOLD, University College London, United Kingdom
CRISTINA BECCHIO, Department of Neurology, University Medical Center Hamburg-Eppendorf, Germany and Italian Institute of Technology, Italy
BENOÎT BARDY, EuroMov Digital Health in Motion, Univ. Montpellier IMT Mines Ales, France
NADIA BIANCHI-BERTHOUZE, University College London, United Kingdom

Movement dataset reviews exist but are limited in coverage, both in terms of size and research discipline. While topic-specific reviews clearly have their merit, it is critical to have a comprehensive overview based on a systematic survey across disciplines. This enables higher visibility of datasets available to the research communities and can foster interdisciplinary collaborations. We present a catalogue of 704 open datasets described by 10 variables that can be valuable to researchers searching for secondary data: name and reference, creation purpose, data type, annotations, source, population groups, ordinal size of people captured simultaneously, URL, motion capture sensor, and funders. The catalogue is available in the supplementary materials. We provide an analysis of the datasets and further review them under the themes of human diversity, ecological validity, and data recorded. The resulting 12-dimension framework can guide researchers in planning the creation of open movement datasets. This work has been the interdisciplinary effort of researchers across affective computing, clinical psychology, disability innovation, ethnomusicology, human-computer interaction, machine learning, music cognition, music computing, and movement neuroscience.

This work was supported by the EU Future and Emerging Technologies Proactive Programme H2020 (Grant No. 824160: EnTimeMent - entimement.dibris.unige.it).
Authors' addresses: T. Olugbade (corresponding author), University College London, Gower Street, London, United Kingdom, WC1E 6BT; email: temitayo.olugbade.13@ucl.ac.uk; M. Bieńkiewicz and B. Bardy, EuroMov Digital Health in Motion, Univ. Montpellier IMT Mines Ales, Montpellier, France; emails: {marta.bienkiewicz, benoit.bardy}@umontpellier.fr; G. Barbareschi, C. Holloway, A. C de C Williams, N. Gold, and N. Bianchi-Berthouze, University College London, London, United Kingdom; emails: {giulia.barbareschi.14, c.holloway, amanda.williams, n.gold, nadia.berthouze}@ucl.ac.uk; V. D'Amato, L. Oneto, and A. Camurri, Universita di Genova, Genoa, Italy; emails: {vincenzo.damato92, luca.oneto}@gmail.com, antonio.camurri@unige.it; M. Björkman, KTH Royal Institute of Technology, Stockholm, Sweden; email: celle@kth.se; P. Keller, Western Sydney University, Sydney, Australia; email: p.keller@westernsydney.edu.au; M. Clayton, Durham University, Durham, United Kingdom; email: martin.clayton@durham.ac.uk; C. Becchio, Department of Neurology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany and Italian Institute of Technology, Genova, Italy; email: cristina.becchio@iit.it.

0360-0300/2022/12-ART126 $15.00
https://doi.org/10.1145/3534970

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Applied computing** → *Law, social and behavioral sciences*; *Health informatics*; • **Human-centered computing** → *Human computer interaction (HCI)*; • **Computing methodologies** → *Machine learning*; • **Social and professional topics** → *User characteristics;*

Additional Key Words and Phrases: Action, databases, datasets, human movement, movement, review, survey

ACM Reference format:
Temitayo Olugbade, Marta Bieńkiewicz, Giulia Barbareschi, Vincenzo D'Amato, Luca Oneto, Antonio Camurri, Catherine Holloway, Mårten Björkman, Peter Keller, Martin Clayton, Amanda C de C Williams, Nicolas Gold, Cristina Becchio, Benoît Bardy, and Nadia Bianchi-Berthouze. 2022. Human Movement Datasets: An Interdisciplinary Scoping Review. *ACM Comput. Surv.* 55, 6, Article 126 (December 2022), 29 pages.
https://doi.org/10.1145/3534970

1 INTRODUCTION

1.1 Rationale

Data is central to the majority of scientific and engineering endeavor [54] with the area of movement analysis and modelling being no exception. The pertinence of datasets that are available for secondary use by the research community is that, on one hand, they represent readily available opportunities to push further the bounds on scientific knowledge and engineering outcomes for relevant areas of research; on the other hand, they also mark the limit of current knowledge and tools, especially where the datasets have been widely (re)used for benchmarking purposes, e.g., for validating machine learning algorithms. While there have been several reviews that cover human movement datasets, they have focused on narrow subgroups, e.g., datasets for human activity recognition, and so only provide very compartmentalized views and critique of the state-of-the-art. Our review sets itself apart with its broader and multidisciplinary survey that enables discussion of contemporary themes crucial to the advance of movement science and technology development.

We focus on human movement primarily because of our interest in it as a means of human interaction with the world (the environment, objects, other humans) and as a modality of expression (of thought, emotion, experience). Beyond our own research investigations, it is obvious from the range of disciplines (e.g., anthropology, arts, cognitive neuroscience, computer science, medical sciences, philosophy, psychology) that cover the topic that human movement is a subject of prevalent interest. The ultimate significance of human movement is that it is fundamental to living and being [74], and attention to it advances bodies of knowledge that could inform, for example, clinical practice or technological development aimed at supporting or augmenting the performance of human activity.

Our scope in the current review excludes datasets that capture a single anatomical location (largely those with face or hand only captured, or those based on the non-optical sensors of a single smartphone) because of the relatively limited information about body movement that they provide. While such datasets are valuable for certain applications, we chose to focus on the very large number of other human movement datasets that include multiple anatomical regions and so are expected to be of wide interest in the research community.

1.2 Aims

One of the primary aims of our survey was to offer a comprehensive list and organised description of human movement datasets that are open for use by the research community. We sought to provide information about the types and contexts of movements captured, the subject population groups, whether the data captured involved single or multiple persons, the sizes of the datasets, the driving research questions, the types and modalities of the data, and how they can be accessed. The main aims of further discussion were to highlight gaps that exist and propose a set of conceptual elements that could guide the creation of datasets that can be more widely reused within and across research communities interested in body movement understanding and analysis.

1.3 Review Approach

In line with the chief aim of our review to present a comprehensive record of human movement datasets, we chose a systematic approach that enables an extensive survey of literature. In particular, we use a scoping review methodology, as it is appropriate for identifying available resources within a given topic [63]. Scoping reviews are further suitable for determining valuable characteristic themes across these resources as well as for uncovering any gaps that exist [63], making them ideal for addressing our aims. We employed the PRISMA-ScR checklist [80], which is an established guide on best practices in a scoping review, to guide our review from conception to write-up.

1.4 A Review of Previous Human Movement Datasets Reviews

In this section, we discuss the 16 previous reviews that we found for human movement datasets, with particular focus on the variety of datasets that these reviews cover, the breadth of information that they provide, and the scope of their discussion.

From Table 1, which gives an overview of the previous reviews, it can be seen that their primary limitation is the limited number and variety of human movement datasets that they cover. In fact, assuming that there are no overlaps in datasets between reviews (which we know is not the case), all 16 reviews would cover only 607 datasets altogether with 96 being the largest covered by any one review. Furthermore, each of the reviews typically only covers a very specific group of human movement datasets. For instance, the reviews of References [20, 28, 29, 41, 43] focused on (RGB) video datasets although [20] included discussion of motion capture datasets as well. Several of the more recent reviews [15, 33, 66, 76, 92] considered the broader RGBD set, i.e., comprising both RGB and depth videos. The review of Reference [33] was not exclusive to human movement but that of References [76, 92] focused on human action/activity recognition in particular and Reference [66]'s was on gait alone. Other reviews looked beyond video data such as the work of References [1, 2] on action recognition datasets in general and the review of hand and arm gesture datasets in Reference [70]. Other examples are References [25, 31, 75] whose reviews were on human action/activity recognition datasets. Only two dataset reviews (References [29, 41]) touch on datasets that can be used for affect modelling (automatic affect recognition [42, 50, 55, 65, 78, 89, 91], in particular).

The merit of the narrow foci of these reviews is the opportunity to give very detailed information about each dataset, and indeed, the most detailed of the reviews captured over 20 variables about each of the datasets surveyed. For several of the reviews [15, 29, 31, 33, 41, 43, 76, 92], analyses of the variables recorded were provided. For example, in Reference [25], descriptive statistics were given for variables including the number of individuals captured simultaneously and sensor categories (wearable, ambient, smartphone). The reviews in References [15, 33] similarly discussed applications of the datasets, e.g., object detection/tracking, semantic labelling. The discussion in Reference [92] also included application domains as well as data formats and levels of complexity of settings and movements. Some of the reviews additionally include detailed discussion of each individual dataset [20, 31, 41, 76, 92] (e.g., highlighting the machine learning approaches explored [76, 92]) or each dataset in a selected subset [25, 75].

However, the reviews are thus limited in the critical gaps that they capture, whereas due to differences in the structure and level of detail across them, they cannot be easily synthesized for an integrated discussion. Our review, however, is based on a survey of datasets from a broader range of

Table 1. Previous Human Movement Datasets Reviews

Review	Year	Focus	Number reviewed	Dataset details included (in addition to name and reference)
[41]	2006	Video datasets including the face and/or body and with affect labels	14	Data type, settings, size, viewpoints, labels, demographics, if acted or spontaneous, anatomical regions, number of modalities, availability
[28]	2010	Sign language video datasets	6	Number of annotated frames, size, application
[2]	2011	Action/activity datasets	37	-
[20]	2013	Labelled video datasets for human action/activity recognition (and motion capture, pose, and gesture datasets)	68	Source, context, settings, studies where used, protocol, purpose, sample images, labels, label format, number of subjects, application, webpage, background type, viewpoints, interaction type, camera movement
[43]	2013	Video datasets for everyday action recognition	18	Number of actions, size, source, modelling performance
[1]	2014	Gesture, action, and activity datasets	23	Annotation type
[70]	2014	Labelled arm/hand gestures datasets	15	Number of citations, size, sampling rate, number of subjects, sensor type and placement, viewpoints, resolution, quality, gesture type and class, label type, anatomical regions, if subject sedentary or not, availability
[33]	2016	Labelled RGBD datasets	96	Sensor, size, labels, data type, camera pose information, number of objects/subjects, settings, application
[92]	2016	Labelled RGBD datasets for action and activity recognition	44	Sample images, size, number of subjects, demographics, data type, settings, viewpoints, context, labels, protocol, annotation type, movement constraints, pre-processing, purpose, background type, details of use
[31]	2016	Human movement datasets	51	Data type, number of viewpoints, webpage, labels, size, number of subjects, application, settings, background type, number of citations, sample images
[15]	2017	RGBD image datasets	46	Number of objects/subjects/events/scenes, size, labels, purpose, target recorded, sensors, number of citations, modalities, camera movement, webpage
[25]	2018	Human activity recognition datasets	5	Citation metrics, modelling approaches, modelling performance, settings, environment, modalities, sensors, duration, number of subjects, activities, labels
[66]	2019	RGBD-based datasets for gait analysis	11	Number of subjects, demographics, sensors, sensor placement, heterogeneity, data type
[76]	2019	RGB(D) datasets of human action	71	Data type, application, studies where used, number of subjects, labels, settings, resolution and sampling rate, size, webpage, purpose, context, activities, background type, number of viewpoints, heterogeneity, modelling approaches, modelling performance
[29]	2019	Affect-labelled video datasets	18	Whether subject, annotator, or contextual information was provided in the dataset
[75]	2020	Human activity recognition datasets	84	Purpose, source, number and type of labels, webpage

Table 2. An Overview of the Search Strategy Used in Our Survey

Database	Search specification	Result size	Shortlisted	Search and shortlist dates
Google Scholar	Title contains: "dataset," "activity database," "action database," "movement database," "motion database," "motion corpus," "movement corpus," "action corpus," "activity corpus," "action data," "activity data," "movement data," "motion data," "motion capture data" (each keyword or key phrase was searched separately, and each search was done by individual year if there were more search results than the database's search cap, 1,000)	28,970	943	21 October–5 November, 2020
ACM Digital Library	Author keyword contains: "database," "corpus," or "dataset" AND Anywhere in the full text includes: "action," "activity," "motion," or "movement"	2,000 (cap)	14	6 October, 2020
IEEE Xplore Digital Library	Author keyword contains: "database," "corpus," or "dataset" AND Anywhere in the full text includes: "action," "activity," "motion," or "movement"	1,228	182	28–29 October, 2020
SpringerLink	Anywhere in the full text includes: "participant" or "human" AND "action," "activity," "movement," "motion," AND "dataset," "database," or "corpus" (search for conference papers and for journal or other articles were done separately)	39,960 (cap = 19,980)	3,464	9 November, 2020–28 January, 2021
PubMed	Medical Subject Heading (MeSH) contains: "activity," "daily living," "movement," "motion," "motor" AND with "Associated Data" ("dataset," "corpus," "action" did not exist as MeSH Terms in the database)	9,814	122	23 November–8 December, 2020
APA PsychInfo	Subject Heading contains: "data*" or "corp*" AND Medical Subject Heading (MeSH) contains: "motion," "movement," "motor," "act*" (search for each of the MeSH terms was done separately)	298	5	10 December, 2020

research areas across biomedical sciences, computing, and psychology. This enabled analyses and discussion to a depth that transcends the individual fields relevant to human movement research.

2 SURVEY METHOD

2.1 Search

We (Author TO) conducted a systematic search of relevant articles via three search engines for scholarly literature (Google Scholar, PubMed, APA PsychInfo) and three publisher repositories (ACM Digital Library, SpringerLink, IEEE Xplore Digital Library) between 21 October, 2020, and 28 January, 2021. These databases were carefully selected for their comprehensive coverage of peer-reviewed research or other technical articles in the pertinent areas of human science and computing. For each database, we tailored our search according to the search functionalities available for the database; but in general, we searched for articles that described human movement data. Table 2 shows the specific search terms used for each database and the number of results returned. In total, there were 82,270 results obtained.

We (Author TO) followed two levels of screening to weed out non-relevant articles. First, we went through all titles and abstracts of the search results and excluded articles that were duplicates (or gave duplicate description of the same dataset), books/theses, patents, citations without full-text available, survey papers, or descriptions of non-human movement data. This resulted in 4,663 relevant articles. At the second level, we read through the full text of each article in this shortlist and further excluded articles that were found to meet the exclusion criteria above (534) or:

- could not be accessed, e.g., due to a pay wall (331),
- were not available in English language (17),
- described datasets based on a single anatomical region, e.g., face only (375),
- presented simulated data, i.e., not captured from real humans (35),
- described still images or far/top view videos (106),
- had unusably limited description of the human movement data (243).

Of the remaining articles, 1,599 of them were found to be secondary references for the datasets that they described. For each of these, we (Author TO) searched for the primary article or website. We obtained a final list of 1,692 datasets (with 278 based on the secondary references found in our systematic search and 34 found completely outside of our systematic search, e.g., from *a priori* knowledge).

2.2 Charting

We (Author TO) charted these 1,692 datasets under 10 main variables that are designed to provide basic information that a researcher could use in determining which of them might be relevant for their work and how to access them and/or further details. These variables further inform our discussion in Sections 3 and 4:

(1) Dataset name and citation - We included these details as information for identification purpose. However, it should be noted that there are instances for which one of the two is unavailable, and further, there are cases in which different datasets have the same name or the same citation.
(2) Purpose of dataset creation - We recorded the purpose for which each dataset was created. Where there seemed to be multiple uses of the dataset in its primary reference, we recorded the most elementary one (based on the descriptions) provided in the relevant publication. For example, a dataset collected for investigation of both human detection and activity recognition would have human detection recorded as the purpose. To enable analysis, we recorded the purposes under themes. A new theme was added to the list if the set of themes collated till that point did not cover the current purpose or if we found prevalence for a new (sub)theme. For datasets based on secondary research data, we specified the purpose as unknown.
(3) The type of data captured (with the number of data instances, the duration of instances, and the number of participants) - We noted the form of the movement data (e.g., video). We additionally noted non-movement data also available in the dataset. We specified the number of data instances and the duration of each instance in seconds. Finally, we specified the number of subjects from whom the data were captured. The last three of these were documented for the purpose of capturing information about the size of the dataset. The number of data instances was not always given; we only included duration information when the number of data instances was available.
(4) Annotation type - We recorded the types of annotation available. When annotations available were not clearly specified in the article associated with a given dataset, we used other information, such as the type of analysis done with the dataset in the publication, to determine what annotations it contains. For example, for datasets used for investigation of automatic person identification, we assumed that the dataset included person identifier annotation.
(5) Data source - With this, we captured how the data collected was obtained. When available and could be summarised in brief, we also noted the setting of the data recording, e.g., sports or walking, naturalistic or acted.

(6) Population group - If the sampling of the participants was not random or based on convenience sampling, then we additionally recorded the population group of focus. In addition to providing contextual information about the movement data in the datasets, this could also enable some analysis of the level of diversity across movement abilities in the datasets available to the research community. However, we did not provide other demographic summaries per dataset, such as ethnicity distribution.
(7) The quantity of people interacting or captured simultaneously - We noted whether data was captured in individual participant settings or in dyads, groups, or crowd settings. Some datasets included more than one of these settings. Where applicable, we highlighted non-specific settings such as surveillance or driving settings. We did not differentiate between settings with multiple people directly interacting, multiple people not interacting, and multiple people within the same space but with only a subset recorded. Finally, we noted settings where a single individual was interacting with a non-human agent (e.g., a robot).
(8) How to access the dataset - Where datasets were available to the research community for secondary use, we typically specified how they could be accessed in the form of a URL either provided in the corresponding publications or discovered by further online search. To check for data availability, we carried out a careful, manual search of the abstract, conclusion, dataset description sections, and footnotes of the associated article. We additionally performed an automatic search for relevant keywords in the text of the article, particularly "available," "access," and "obtain." For named datasets, we further searched the Internet (using Google as the search engine) for websites or other resources with information about how the dataset could be accessed.

    Although we (Author TO) accessed each URL ourselves and noted cases where URLs provided by the authors were no longer valid or had been repurposed, we cannot guarantee that the URLs enable access to the respective dataset and do not lead to unsafe websites, as they may have become invalid or repurposed since our charting. Thus, we cannot be held responsible for any damage or distress caused by following the URLs noted in our charts or any other use of our dataset catalogue. We urge users of the catalogue to take the highest precautions in their use of the URLs. Finally, we could not always ascertain whether the data available included raw data or comprised extracted features only.
(9) Motion capture sensor - For datasets that include motion capture data, we recorded the type of motion capture sensor used. We did not usually note the sensor brand; an exception is the Microsoft Kinect sensor, which we specified where applicable given the widespread use of this sensor compared to other vision-based markerless motion capture systems (we expected this based on our own use and what we know from interactions with other researchers; our expectations were confirmed by our findings in Section 3).
(10) Funding - For this, we relied solely on the funder specified in the acknowledgement sections of the associated publication. We excluded funders that only provided computing resources, e.g., GPUs. For datasets based on secondary research data, we documented the funder as unknown.

In line with the aim of our survey, the rest of this article largely focuses on the 704 datasets that we could ascertain are open to the research community for secondary use.

3 SURVEY RESULTS AND ANALYSIS

The primary outcome of our survey is a catalogue (see the main document in the supplementary material) of 704 human movement datasets that are open for secondary use within the research community. For each, we provide information (see the descriptive variables in Section 2.2) that can be useful to researchers in finding datasets relevant to their interest. We also include an abridged documentation of the datasets, i.e., the corresponding references and year periods only for an overview of the catalogue, with the supplementary material.

Although we did not limit the time scope of our search (i.e., we did not put any bounds on the years that the search should cover), the open datasets that we found fell into the period between 1997 and early 2021, both inclusive—this range does not include datasets whose creation year we were unable to verify. See the abridged document in the supplementary material for an overview of the distribution of the datasets across year periods. Of the 704 datasets, 38 were only available on request to the authors (e.g., via email); the dataset webpage provided by the authors (or a secondary reference) were obsolete for 79 of the datasets; and how others could access the data was not specified for 95 datasets. Ideally, the process of requesting access should be clear, straightforward, and enduring (and so not dependent on changes to a corresponding author's affiliation, for example). The above findings begin to highlight some of the challenges inherent to the sharing of data within the research community. We point out additional difficulties in Section 5 based on barriers that we encountered in finding (open) human movement datasets for our survey.

In the rest of this section, we provide further findings based on analyses by themes that emerged from the data, guided by the aims of the survey.

3.1 The Main Drivers

In this section, we give an overview of the funding sources for the datasets reviewed and the purposes for which they were created to provide insight into factors that have been responsible for the growth of (open) human movement datasets. We analyze the two variables (funder and purpose) independently. However, we may expect that there is a relationship between them especially at the low level of specific funds (e.g., types of grants), but perhaps also at high-level categories (e.g., public versus industry versus other private funds).

3.1.1 Funders. In our review, there were 469 datasets that had their funding sources specified. We grouped each of the funding sources into one of four categories that emerged from the data: (i) publicly funded or governmental organisations, (ii) publicly funded multinational consortia or unions (made up of n>3 countries), (iii) private business enterprises or their subsidiaries, and (iv) other privately funded institutions. We used information provided in the corresponding articles and careful Internet search to resolve the categories for each of the funding sources. There were only 24 sources for which we were unable to determine the categories. The findings of our analysis of the funding sources for the 469 datasets are reported below.

Figure 1 shows the distribution of the datasets across the four funding categories. It should be noted that several of the datasets had multiple funders; 11% of the datasets had industry (i.e., private business) funders, while 9% had other private sources of fund; 24% received funding from a multinational consortium/union of which the European Union was the primary funder with n = 113 datasets out of the 114 in that category; 71% of the datasets benefited from funding from more local public institutions or government agencies, or grants funded by these. Figure 2 shows a chart of the 40 different countries represented by these 71%. The United States funded the highest number (n = 88) of datasets, followed by China (n = 49), Germany (n = 29), Spain (n = 20), and the United Kingdom (n = 18) in the top five.

3.1.2 Creation Purposes. We further grouped the datasets in our survey with respect to the purpose for which they were created. Multiple purposes were specified for some datasets. As mentioned in Section 2.2, for these, we recorded only the primary purpose specified. There were a number of datasets (largely secondary datasets) for which we were unable to determine the

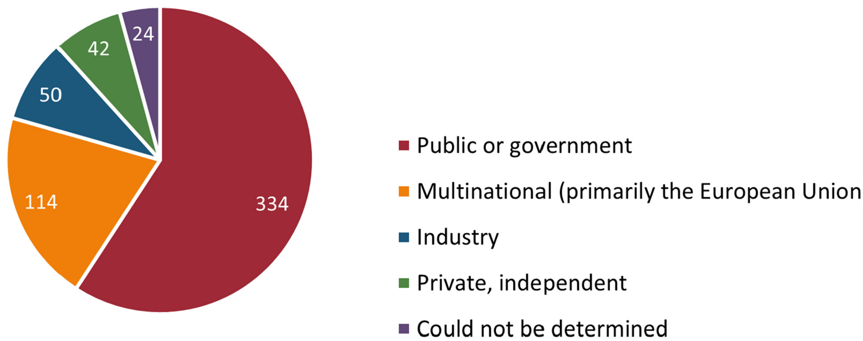

Fig. 1. The distribution of funding categories for the datasets.

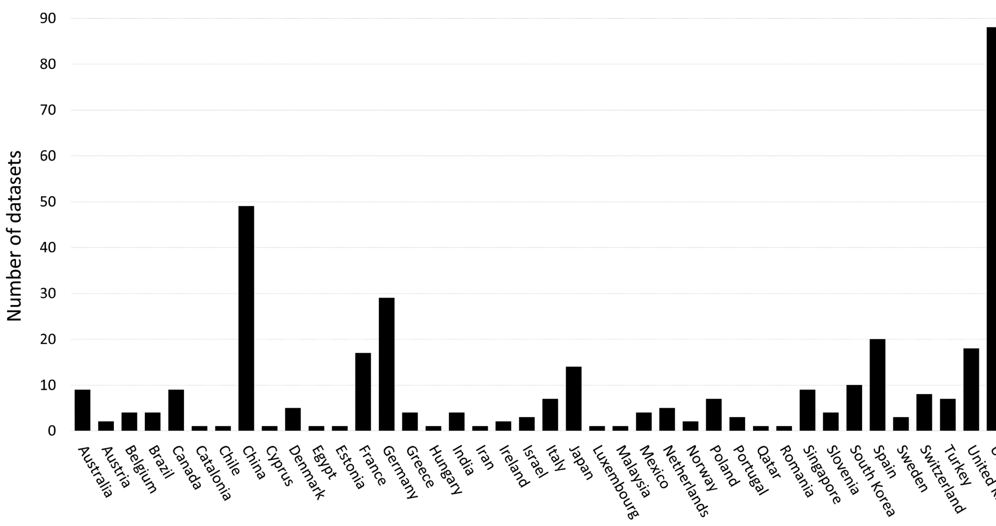

Fig. 2. Funding sources by country for the datasets with public or governmental funding, arranged in alphabetical order.

creation purpose. We classified the purposes we recorded for each dataset into 27 broad categories shown in Figure 3.

As can be seen in the figure, action or activity modelling (or other analysis) was the single most prevalent primary use of the human movement datasets with n = 182. Similar purposes included event, behaviour, interaction, group, and crowd analyses, which together made up the next largest purpose (n = 103). Other related uses were gesture analysis/modelling of which nearly half (n = 19) focused specifically on sign-language gestures. A few other specific movement types bear mention as foci of movement dataset creation. One of these is gait, which had considerable interest in its use for biometrics (n = 22), i.e., human identification, although it was additionally used in other contexts (n = 22), e.g., clinical analysis of gait. Falls are another of these specific movement types with n = 21. With only n = 3, violent movements or fights are one more specific movement type that was focused on.

The second single most prevalent purpose was for affective computing or other emotion studies (n = 50). Other purposes related to this were: movement or intention prediction, modelling of skill

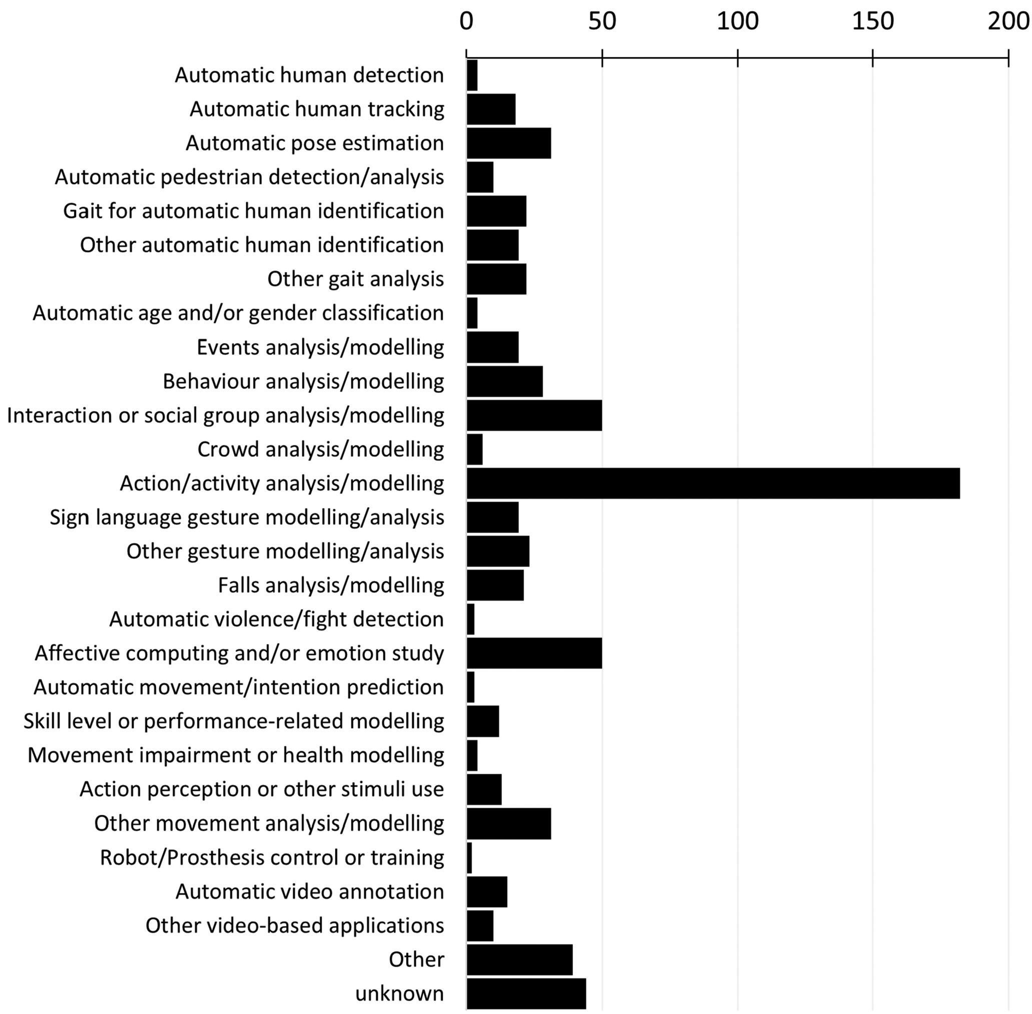

Fig. 3. Frequency distribution of creation purposes across the datasets, with the purposes ordered such that similar purposes are arranged closer together than disparate purposes.

level or other performance measures, modelling of a movement impairment or health condition, and as experiment stimuli. All of these made up 32 datasets. Fundamental movement computing needs, i.e., detection, tracking, and pose estimation, together covered a larger number of datasets (n = 63).

We additionally analysed the recorded purposes to see what research communities were apparent from the data. Although there were overlaps and blurred separations between distinct areas for some purposes, nine main areas of research were salient (reported here alphabetically):

(1) Affective computing
(2) Animation and related
(3) Biometrics
(4) Clinical research
(5) Computer vision

Fig. 4. Movement contexts represented in the datasets (for keywords that occur in at least three datasets).

(6) Human activity recognition
(7) Psychology and social science
(8) Robotics
(9) Other areas of computer science

3.2 Data Types, Sources, and Settings

This section presents a description of the types of movement covered by the datasets, how they were acquired, and the forms in which they were captured.

3.2.1 Movement Settings. We were able to deduce the settings of 539 of the datasets. The majority (n = 294) of this subset comprised data captured in individual settings alone. Still, several datasets covered other settings. For example, 48 datasets represented dyads, 53 had a composition of groups only or both groups and dyads, 65 were made up of other combinations, 16 were based on crowds alone, and 50 were captured in non-specific settings particularly in traffic or surveillance.

3.2.2 Movement Contexts. Where we had the data available (n = 413), we further analysed the contexts of movement in which the datasets were acquired to understand which have been the most prevalent in the area of movement analysis/modelling. Figure 4 provides a visual representation of the contexts that we found to exist in at least three datasets (thus showing only 34 out of 57 contexts in total). The figure shows the prominence (in terms of occurrence frequency) of each context keyword using bolder and bigger fonts for the most frequent keywords. It can be seen that walking has had the greatest interest (n = 70), with sports being second (n = 47), and everyday movements (n = 39), e.g., in home settings, third. Other popular movement contexts (n ≥ 20) were: surveillance, exercise-related movement, conversation, sign-language gesturing, and other forms of gesturing.

3.2.3 Data Sources. Most (n = 549) of the datasets captured in our survey were created via direct recording by the researchers. This has clearly been the traditional means of obtaining human movement data. Given the difficulty of capturing a large number and variety of spontaneous movements in real (as opposed to staged) activities, it is not surprising that other data sources—not particularly any less staged than those recorded by researchers—such as movies (n = 14), TV broadcasts (n = 28), YouTube (n = 35), other internet sources (n = 21), and crowdsourcing (n = 3) have been explored. As one would expect, while sourcing of data from movies and TV went as far back as 2004 in our survey, the first occurrence of data acquired via any of the other methods was slightly more recent (i.e., 2007). Sports was one of the largest represented movement contexts captured in the datasets based on either YouTube (n = 10) or TV broadcasts (n = 9). A number (n = 34) of other datasets have also been built from existing research datasets. For example, the AVA dataset [40], which was originally developed for use in automatic action recognition investigations, was repurposed as the BoLD dataset [57] with further annotations for affective computing use.

3.2.4 In the Wild (or Not). Most of the datasets comprised data that are best characterised as belonging somewhere in the spectrum between definitely acted (n = 36) and clearly naturalistic (n = 39). However, a number of datasets included data captured in the wild (n = 36), i.e., in organic settings that, unlike merely naturalistic settings, are not purposefully recreated for the goal of collecting data.

It should be noted that for data acquired from the internet (including the YouTube platform), especially those that cover everyday movements, it was not usually clear to us whether they represented in-the-wild settings or if they included staged activities.

3.2.5 Data Type. As can be seen in Figure 5, with a few minor exceptions (muscle activity data (n = 16) and localization data (n = 1)), there were three main formats of movement data: video, joints positions or angles, and **inertia measurement unit (IMU)** data.

Video was the most predominant form of movement data with 53% of the datasets based on video data only. Of these, 93.5% included (or were based exclusively on) RGB video. This finding is not unexpected, as although video cameras (especially RGB) are highly privacy intrusive and can be used for covert data capture, they are a convenient (and cheap) means of capturing body movement. We should remark here that movement data acquired via YouTube, TV, and movies (as opposed to those recorded by researchers themselves) are confined to videos, although recent advances in computer vision, e.g., OpenPose [17], have made it possible to extract kinematic features from video data. It should also be noted that the limited confidentiality offered by video formats often mean that captured videos are not released in open datasets that include them. In such cases, features extracted from the videos or other available kinematic data are instead made open for reuse.

Beyond RGB videos, depth videos were another popular form of data capture, perhaps facilitated by the Microsoft Kinect sensor that became a readily accessible system for acquiring RGB and depth data simultaneously (together with joints position data, as will be discussed below). There were 145 datasets that featured data collected via depth video only or in combination with other types of video data; 101 of these datasets were an exclusive combination of RGB and depth video. Some other video data types that we found in our survey include: grayscale or monochrome, thermographic (i.e., thermal, shortwave infrared, or other infrared), point-cloud, lidar, point-light, silhouette.

Only 34% of the datasets in our survey included kinematic data in form of joints positions or angles; the majority of these comprised full-body joints data. We further analysed the sensor types used to collect these joints position/angle data, and we found that Microsoft Kinect (n = 109) and marker-based optical motion capture systems (n = 79) were the most commonly used sensors. Even

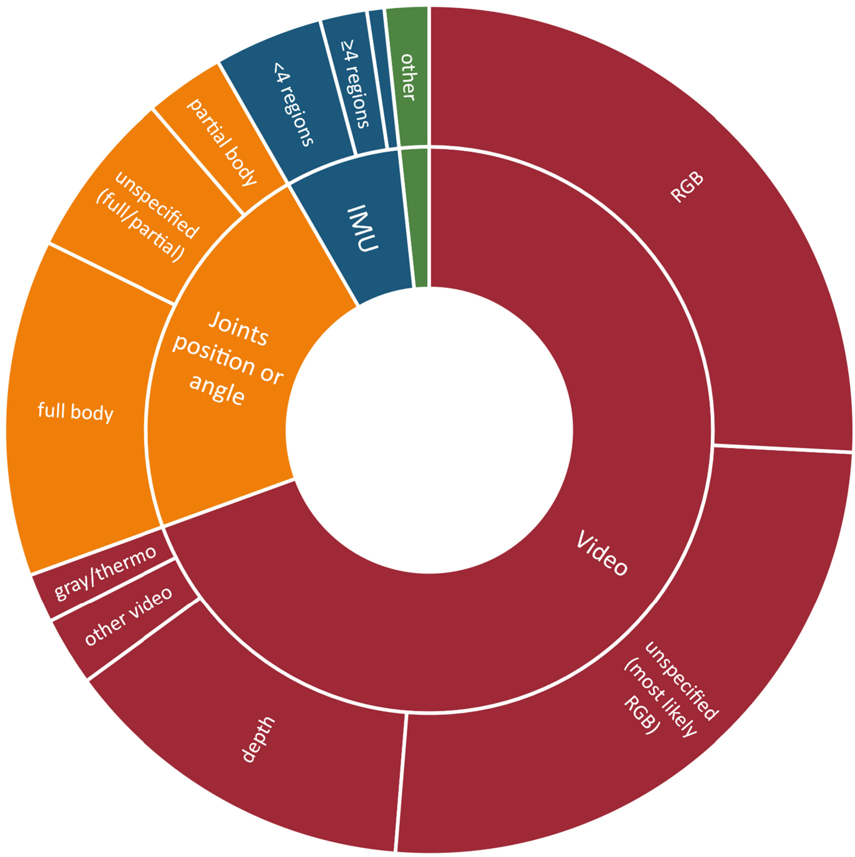

Fig. 5. The distribution of movement data types in the datasets (the "other" slice represents the minority muscle activity and localization data types). Note that the "gray/thermo" slice represents grayscale (or monochrome) / thermographic (including thermal, shortwave infrared, other infrared) videos.

less striking in number (possibly due to our exclusion of datasets with data from a single anatomical location only in our survey) were datasets that comprise data captured from IMUs (accelerometers, gyroscopes, magnetometers). There were 70 datasets (10%) that included this form of movement data.

Some datasets included additional data beyond movement data. For example, for 50 datasets, audio data was explicitly captured. Similarly, 14 datasets included ground force reaction data, 8 included gaze data, and 4 included GPS data. Other data types included were: ambient data (e.g., ambient temperature), interaction data (such as motion sensor, tilt switch data), and physiological data (for example, respiration, electrodermal activity).

3.2.6 Annotations. On analysis of the types of annotations provided for each of the datasets, we found that they fell into seven main categories:

(1) Movement type (Movt) - e.g., action or activity labels,
(2) Auxiliary movement description (AuxMovt) - such as types of walks or compensatory movements,
(3) Affect - including cognitive states, experience, personality, and other traits,

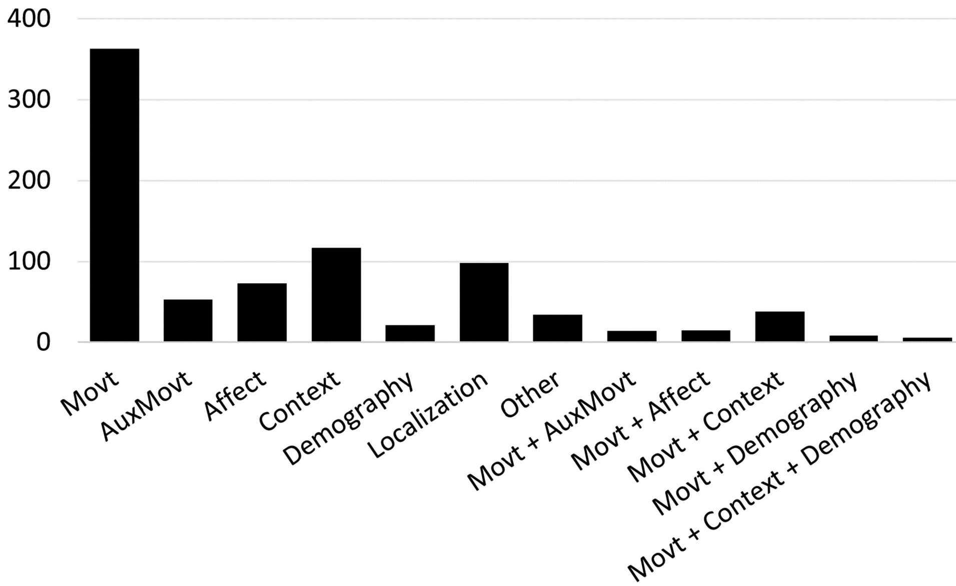

Fig. 6. Occurrence frequency of each of the seven annotation categories. Movt: movement type, AuxMovt: auxiliary movement description.

(4) Contextual information (Context) - which could be valuable in interpreting corresponding movement data, for example, speech or other acoustic labels, health information, person identifier, the types of objects being interacted with,
(5) Demographic information (Demography) - e.g., information about age, sex/gender, height, skin colour,
(6) Localization - only applicable to video data and usually based on the use of bounding boxes, it covers localization of specific body regions (such as the head), whole bodies, or an activity or interaction,
(7) Other - any other labels that did not belong to the above classes.

As shown in Figure 6, movement type annotations were the most occurring forms of labelling, n = 363. Contextual information and localization information were the second and third most frequent forms, although they occurred far less than movement type labels with n = 117 and n = 98, respectively. There were even much fewer occurrences (n = 38) of combinations of movement type and contextual information annotations. Affect-labelled datasets were similarly a minority (n = 73), more so those that additionally included movement type labels (n = 15).

3.3 Population Groups

We looked at the categories that emerged with respect to the population groups covered in the datasets. Figure 7 gives an overview of the five main categories we found. Each of these categories further belong to one of two classes: general or specific population groups.

There were four main categories under the specific population groups. The predominant of these were experts in movement or movement-related expressions, particularly art performers (i.e., actors or dancers, n = 49), athletes (n = 39), and signers (n = 11). It should be noted that professional actors were usually used as data subjects not because of specific interest in them as a group, but rather for their proficiency in providing non-functional movement performance or expressions on

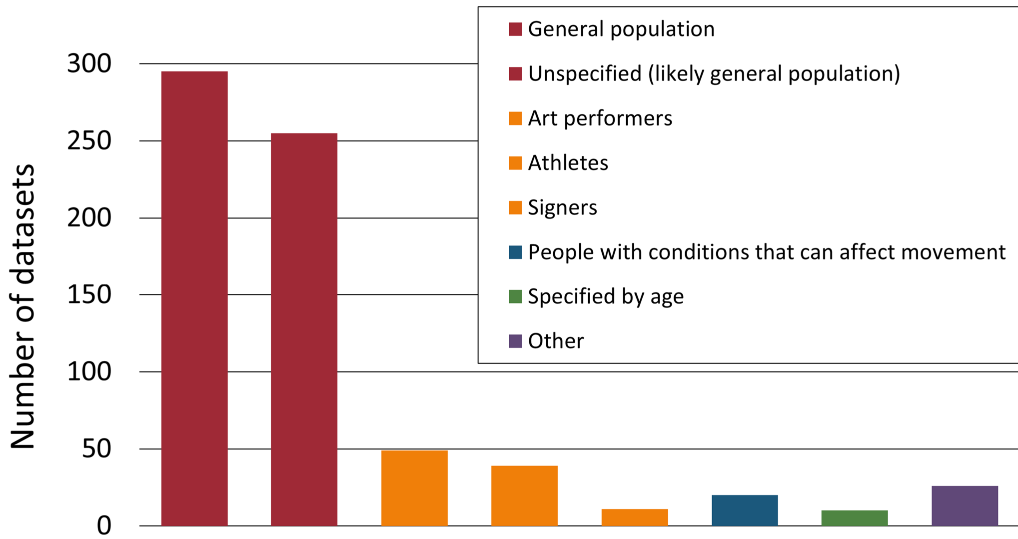

Fig. 7. The five main categories of population groups of interest in the datasets, coded by colour: "red": general population; "yellow" : experts in movement or movement related expression; "blue" : people with medical conditions that can affect movement; "green" : specific age groups beyond adults; "purple" : others.

demand. Although limited in number (n = 20), another category of specific groups that we found were people who had medical conditions that could affect movement, e.g., Parkinson's disease, limb loss, stroke, chronic pain. There was also a category defined by specific age groups other than adults, which are those usually represented in the general population class. For instance, six datasets included data acquired from children and four included data collected from the elderly. The fourth category covers other specific population groups that did not fall into any of the other three categories, for example, Reference [85] collected data from automobile factory workers, and Reference [61] obtained data from firefighters.

The majority (n = 295) of datasets contained data from the general population. Data subjects were usually sampled using convenience strategies, e.g., recruiting colleagues, students, or anyone available who responded to the recruitment ads. Due to ethical constraints and data protection laws, they were typically adults at the time of the data capture especially for data collected by researchers themselves, as opposed to secondary datasets such as those sourced from the Internet or movies. There were additional datasets (n = 255) for which we were unable to determine which of the five categories their data subjects belonged to. Most of these are likely based on data from the general population.

4 DISCUSSION

Given the findings in Section 3, we revisit the state-of-the-art in human movement datasets on the basis of four main considerations: human diversity, ecological validity, the multifacetedness of movement, and ethical issues with implications for secondary use. Through this discussion, 12 dimensions of human movement datasets emerge, culminating in the framework in Table 3, which is intended as a tool to guide the creation of datasets, especially where authors intend to make them open to research community use. The framework aims to be a starting point for researchers and to enable them to think beyond their own primary use of the dataset to maximize reuse. This is ultimately critical for cross-fertilization between studies and disciplines, cross-validation of findings,

Table 3. A Framework for Human Movement Datasets Creation

	Dimension	Description	Rationale
Human diversity	Body configuration & neurodiversity	Body morphologies and movement-related (dis)abilities. Variations in social, learning, and other psychological functions.	Current datasets mainly represent certain group of peoples and in limited (stigmatizing) contexts for some of those. This limits understanding of movement and constrains the value of clinical practice or technology that they inform.
	Culture & geography	Cultural context, geographical region, and ethnicity	
	Age & others	Age groups, genders, education levels.	
Ecological validity	Spontaneity	The spectrum between movement performed on cue (e.g., acted) and movement initiated to achieve a goal completely outside of a research or other study	Movement is affected by the circumstances from which it emerges and the settings in which it takes place. While some studies may benefit from fully controlled protocols, others may require naturalistic settings, i.e., close to real-life situation, or capture in the wild where attainable.
	Environment	The physical space (including configuration, mobility, environmental conditions, objects, familiarity, space ownership, i.e., research vs. personal vs. shared vs. public space)	
	Social context	The presence of others (i.e., in small/large groups) and the type/level of interactions	
	Psychological experience	The state of mind (e.g., intentions, emotions, competing goals, motivations, desires)	
Movement as complex phenomena	Movement sensors	Sensors to capture observable body movement (e.g., cameras, IMUs, motion capture)	Various processes (including psychological, neural, physiological) are involved in movement and as such they should be considered when deciding movement data to record. They could help provide a more in-depth understanding of movement, its triggers, and its significance and implication.
	Para-movement signals	Physiological (e.g., muscle activity, respiration, brain signals) and other behavioural signals (such as gaze) concomitant with and/or critical to movement execution	
	Movement description	Levels of movement abstraction: pose, low-level features (e.g., flow), action (or gesture, behaviour, other event), interaction type or level, activity, affective/cognitive qualities	
	Contextual information	Additional data useful to account for in modelling movement	
Ethics & data protection		Ethical issues (e.g., consent, cultural or potentially sensitive issues) in the collection and processing of the original data. Data protection regime under which data was collected.	Limited information on ethics and data protection for the original data can make secondary use a challenge in countries where such information needs to be checked against local regulations of the country of reuse.

as well as resource sustainability. We further call for a *Human Movement Bank*, i.e., a common, multidisciplinary human movement repository where datasets can be shared and more easily discovered (see Section 5 and the introduction in Section 3 for challenges currently faced in finding and accessing open movement datasets). Beyond its value to individual research groups, such a repository has the potential of fostering interdisciplinary collaborations and additionally being a space for researchers to contribute data collection protocol templates and tutorials that can be of value to other researchers and facilitate good practice in the research community.

4.1 Human Diversity

Our findings in Sections 3.1.1 and 3.3 highlight two gaps in relation to diversity in the collection of existing open datasets. One is linked to the underrepresentation of diversity of both minds and bodies. Few, for example, include people with disabilities and neurodiverse individuals beyond the scope of medical questions such as diagnosis, rehabilitation monitoring, or clinical analysis. We found no datasets including people with disabilities performing sports, engaging in artistic expressions, or simply performing everyday tasks. One exception is Reference [24], which features movement data from children with autism spectrum disorders and was possibly captured in the wild, but the purpose of this dataset was the identification of repetitive and self-stimulatory behaviours, a practice that has been heavily criticised by the autistic community as disability surveillance associated with oppressive practices towards individuals rather than promoting more inclusive society [88]. As highlighted by both References [77, 87], this skewed representation of people with disabilities has negative implications both from an inclusion perspective and from a technology one. It additionally perpetrates stigmatizing views of people with disabilities, who are restricted to the role of patients in need of medical help, rather than shown as agents according to more progressive and socially aware models of disability. Further, it leads to incorrect assumptions that shape the design and development of new technologies to automatically exclude people with disabilities from being considered as potential users [18]. The lack of inclusion of individuals with disabilities has been documented by numerous authors in the field [37, 77, 87] although from the perspective of technological systems (e.g., AI, virtual reality) rather than human movement datasets. Future datasets for movement analysis should pursue inclusivity. For instance, including people with missing limbs or minds that process information differently from the neurotypical can enable the expansion of the conceptualization of the human body towards plurality and diversity outside of normativist tendencies [27]. This can be critical, e.g., for technology designs that are inclusive and do not exacerbate exclusion and stigmatization. For studies focusing specifically on the body movements of people with disabilities, dataset creators should consider settings and purposes beyond the clinical, incorporating the diversity of activities that are part of their full life experience. While it may seem idealistic to include people with disabilities in datasets for applications relevant to the general population or make data capture specific to a given disability much more focused on the everyday challenges for people with this disability (rather than only on clinical or lab settings), there are considerable opportunities to do so. For example, the latter should be underpinned by maximal involvement of potential participants from the relevant population groups in setting research questions, contributing to research design to address those questions and helping researchers to be as inclusive as possible in their recruitment. Indeed, there are policies, e.g., in the UK, that make this a requirement, with funding schemes made available to encourage it [56]. These principles of **patient and public involvement (PPI)** (see Reference [48]) are only very partially realized in the datasets found in our survey.

The other gap in term of diversity is the limited representation of geographical regions, which can be roughly deduced from Figure 2. Ethnicity (both in diaspora and native communities) and culture (beyond ethnic affiliation) are blueprints for the way people move, express intention, and show affective qualities [79]. There are also differences in the types of activities (e.g., sports, leisure, occupational) more commonly observed and how they are performed across cultures and geographies [9, 46]. Moreover, ethnicity has further been linked to differences in body morphology that can influence movement [86]. While there is some level of cultural and geographical diversity across existing open movement datasets particularly for musical and dance performances (such as the IEMP corpus [23, 49, 68] and the AniAge project [3]), the extent of recorded scientific knowledge on movement is largely based on data from (as well as on researchers in) North America, Europe, Australia, and parts of Asia, as shown in Figure 2. This is similar to findings in Reference [45] that a significant number of psychology studies were based on a limited representation of contemporary human societies. Worthwhile goals for future research should include the creation of datasets that capture other geographical regions (not to the exclusion of diversity of other characteristics such as gender and literacy backgrounds). To enable comparison across cultural/geographical contexts, model datasets could be reproduced for novel cultural or geographical contexts. Such endeavours need to be integrated in collaboration with experts and participants within the relevant communities rather than merely considering them as data subjects (or annotators) [5, 71]. Particular care needs to be taken to not approach the work in a way that is or could appear exploitative, in addition to other ethical considerations.

Our proposed framework in Table 3 features three main dimensions that touch on these themes of inclusion and diversity. In addition to the elements of body configuration, neurodiversity, culture, and geography discussed above, we include age and other significant characteristics, such as gender and education, as an additional dimension. Our findings in Section 3.3 suggest limited coverage of age groups outside young adults although age is well known to affect human movement [90]. A possible reason for the limited availability of datasets of children could be the ethical sensitivity of the data, making it challenging to open it to the wider research community especially when it includes personally identifiable data such as videos. While some of the general datasets (i.e., not age-specific) represent a good spread across adult age groups, it is nevertheless necessary to emphasize the importance for age groups to be well-represented to ensure that our understanding of movement as a community reflects the population. The same holds true for gender and literacy and education levels. While recruitment practicalities can be a limiting factor, it is important to consider the merits and demerits of various sampling techniques in terms of how they could constrain the distribution of age and gender groups and education levels [32, 81]. Endeavours to address diversity could include creation of datasets that focus on groups underrepresented in existing datasets. This may be more attainable than representation of multiple groups in a single dataset. Although our framework only highlights diversity in terms of body configuration, neurodiversity, culture, geography, age, gender, and education levels, for certain applications it may be valuable to consider other relevant dimensions. Taking the case where conversation analysis is of primary interest as an example, it may be worthwhile, for instance, to consider actively recruiting people who employ other means of communication beyond speech (perhaps due to speech impairments), e.g., sign language, alongside those who use speech. Beyond enriching the understanding of conversation, as well as being inclusive, this additional dimension can expand the secondary use of such data to the areas of sign-language gesture modelling, for example. Finally, we clarify that rather than expecting all datasets to be all rounded, our aim is to: (1) guide dataset creators in thinking about human diversity when planning movement data collection; (2) encourage collaborations that seek to expand available datasets by better capturing underrepresented groups, rather than merely reproducing any skews that exist; and (3) provoke a discussion across movement research communities on the extent to which scientific knowledge and developed technologies represent the world population as well as plans toward a more comprehensive understanding of movement.

4.2 Ecological Validity

Findings in Sections 3.2.1 and 3.2.4 further point to limitations in the naturalness of activities in most existing open movement datasets. First, many datasets were captured in staged settings (acted/instructed, or naturalistic, but not in the wild). Although capture in acted settings (e.g., movies) enables exploration of movement expressions that may not readily occur in real life (such as violent behaviour), acted expressions are based on stereotypes and often lack the subtlety of natural expressions [55]. Similarly, datasets where movements are performed on instruction in lab settings (e.g., walking, reaching, and grasping) miss out the medley of factors that have real implications and influence movement in everyday life but are difficult to recreate outside the real situations in which they occur. For example, walking could be in the context of walking in a busy train station and anxious to catch one's train, and reaching and grasping could be in the context of deliberating on the price and expiry date on several items during grocery shopping before choosing one [53]. More naturalistic movements in real activities (e.g., dance or music performance, cooking) address some of these problems, but in constrained settings or otherwise outside of in-the-wild situations, they are also not fully representative of real life behaviour. While controlled studies are valuable for studying movement, it is important to additionally investigate experience and behaviour in context to further inform the development of practices and technologies that support real human activity.

Second, although our findings show many datasets recorded in social settings, a large number do not capture the social interactions in which the movement of interest occurs, rather than obtaining data from the individual in isolation. Authenticity in social context is critical for more representative models of individual movement and related expressions, since movement is constrained by the presence of others, and the propagation of intention or emotion components in human groups influences both internal regulation and environment exploration [10]. For instance, affective or artistic expression in musical or dance performance may unfold differently when there is an audience [62]. Recording social interaction of course becomes necessary for modelling relational and group activities and emotions (e.g., team dynamics and leadership in sports and music/dance performance, dance, human-robot interactions, and conversation). The relative paucity of (open) data on social scenarios in our survey highlights the domination of "first-person." behavioural sciences (neurosciences, cognitive sciences, or biomechanics) [14, 73].

We highlight four dimensions in our framework (Table 3) that cover the level of ecological validity of recorded movement: spontaneity (how much the movement is acted, elicited, or spontaneous); environment (how much the physical space represents in-the-wild settings); social context (the type and level of social interaction involved); and psychological experience (the affective and cognitive states that drive how movement is executed in the real world). Using the example of body movements associated with musical performance, which is an active and growing area of research [22], rather than simply recording a lone musician in a lab space that may be sufficient for the immediate needs of the primary data user, it may be globally worthwhile to explore the possibility of incorporating relevant social interaction. Whether that is other performers, an audience, and/or dancers would depend on the primary research question(s). Emotional experiences (which can in turn depend on the social context, e.g., excited audience paying to see a performance vs. researcher-picked cohort paid to take part in a study) could additionally influence the representativeness of the performance behaviour of the musician of interest.

4.3 Movement as Complex Phenomena

Our findings in Sections 3.2.5 and 3.2.6 highlight four movement data dimensions that are critical to consider in creating a movement dataset. We discuss each of them below. We additionally call attention to one of the opportunities for which open movement datasets could be leveraged to manage some of the challenges in acquiring movement data covered all through the discussion section.

4.3.1 Movement Sensors. Our findings in Section 3.2.5 suggest a preference for the use of video cameras to capture movement data although motion capture sensors from which high-fidelity, 3D kinematic data can be extracted was favoured in a good number of studies. Among other considerations, movement sensors can further influence ecological practicalities, as researchers often have to balance data needs with logistical constraints. For example, video data suffers when there are occlusions, e.g., furniture in the home, other performers in a live dance/music performance. Occlusions usually mean loss of data [4, 52], although they can themselves be data of interest such as for fall detection in Reference [26]. Videos may not be practical when a high level of mobility is involved, e.g., doing laundry (which can involve multiple rooms in the home), shopping (which can involve multiple floors in the same building or different buildings). Other favoured sensor systems, particularly markerless motion capture and marker-based optical systems, have the same limitations. IMU-based wearable sensors is an alternative but not without drawbacks; for instance, the person(s) being captured needs to (remember to) wear the sensors.

A more promising direction for practical recording of movement data in real unconstrained settings may be hybrid tracking, for example, combining vision-based sensors (e.g., markerless motion capture) with wearable IMU sensors [19]. Another possibility, although further in the future, is the development of mobile motion capture equipment (perhaps drone-like) that does not interfere with the tracked activity or interaction, has capacity for constrained physical spaces such as in the home, and allows in-parallel zooming in and out of specific body regions of interest such as the ankle of a runner, the body-hands-head synchronization of a kayaker (for instance, the EuroMov flying carpet sensor system). Emerging technologies such as impulse radio ultra-wideband sensors [12] could further expand the number of movement sensor options available to dataset creators.

4.3.2 Beyond Body Movement Data: Para-movement Signals. Multimodality is valuable, perhaps even vital in certain use cases, to movement modelling given the different (behavioural, neural, physiological) layers of movement execution [6]. For example, brain activity can provide insight into action planning mechanisms and strategies that occur at multiple timescales [13] and explain interpersonal coordination in joint action [51]. This is unsurprising given that the brain is involved in motor control, processing of spatial and temporal information, and motor learning [36]. Another process relevant to movement is respiration, which has, for instance, been found informative for capturing differences in movement qualities [58]. Respiration has further been implicated in motor control itself, and it is additionally associated with emotional experiences that can influence movement [83]. Multimodal datasets can thus provide insight into the internal milieu of body movement beyond what is observable. Such insight can, for instance, inform analysis of the dynamics of socio-motor interaction with other agents or surroundings by comparing it to the interoceptive inference of the agent via qualitative self-report [7].

However, findings in Section 3.2.5 show that only a few datasets featured data beyond body movement. This is possibly due to a limited understanding of movement as complex multidimensional phenomena in certain research areas. As such, our framework aims to highlight the opportunity provided by a multimodal approach to movement modelling. At the same time, there could be difficulty in capturing multimodal data due to increase in logistical complexity and further constraints on ecological validity resulting from sensor demands. For example, while the capture of muscle activity, another modality directly relevant to movement [67], can be fairly unproblematic in lab settings, sensor costs, compactness, and attachment convenience can make it a challenge in home studies where the participant has to attach the sensors themselves (compact and convenient sensors can be more expensive than is practical for such types of studies). Among other directions, research investigations into the possibility of extracting various physiological signals from data available from movement sensors (e.g., respiration from low-cost thermal video data [21]) can expand the feasibility of multimodal data capture.

4.3.3 Movement Description: Levels of Movement Abstraction. Findings in Section 3.2.6 suggest that annotations have typically only covered either the traditional levels of movement abstraction (i.e., postures, actions or gestures or behaviours or other movement events, interactions, activities) [31] or affective and cognitive experiences that can be higher levels of interpretation of movement, but not both. Meanwhile, both behaviour categorization (the what) and semantic interpretation (the why) are encoded in human action perception, with each playing a role in interpretation of observed behaviour as well as in responding appropriately based on previous experiences [34, 47].

Including labels for as many levels of abstraction as possible can be valuable in advancing the state-of-the-art in many areas of research, including development of automatic detection models, building of behaviour generation models for artificial agents, clinical analysis of movement, and neuroscientific modeling of socio-motor interactions. For example, emotions are not only interesting from the point of their manifestations in movement, but how they change the predictions that agents make with regards to their environments and therefore their subsequent behaviour [8] (i.e., think about how different a person is going to behave and feel walking into a room of people expecting a threat versus a room of people being cheerfully engaged in a social gathering).

Thus, multilevel annotation of movement data including both behaviour and psychological state is useful. However, annotation is expensive, and each additional annotation level adds to challenge. One practical solution employed for a few datasets is to add new layers of annotation to existing labelled datasets, e.g., adding low-level behaviour labels to movement datasets that already have affective labels.

4.3.4 Leveraging Existing Datasets to Manage Limited Data Sizes. As highlighted in previous subsections, acquiring movement data for certain settings or population groups can be challenging, and it may be possible to obtain only a limited data size for such contexts. Following approaches typically used in other research areas, deep neural network models pre-trained on existing movement datasets could be used to extract valuable movement representations (encodings, embeddings, deeply learned features) from raw sensor data in the newer dataset, such as the use of the pre-trained VGGFace2 [16] in face modelling and the pre-trained ResNet-50 [44] for computer vision tasks, e.g., image recognition. Pre-trained models can be particularly useful for obtaining low- or mid-level features that would usually require deep neural networks that depend on very large data sizes. They can also be useful for reducing movement data dimensionality for computational modelling or other analysis. This highlights value in extending available good quality datasets (i.e., datasets obtained following best practices) in the area of movement analysis to include additional annotations, population groups, modalities, and contexts, as such expanded datasets can enable the development of pre-trained models that cover a larger variety of movement and are as well more relevant to downstream applications, i.e., secondary uses of such models. In other similar approaches, learned representation from one or more source settings can be transferred to a new target setting that has some commonality with the source setting(s) [69, 93]. For example, one could assume that for the same modality, similar movements across different settings will have a common feature space, and so the learned encodings from one setting can be used to regularise the learning of encodings for a different application context. Of course, care would need to be taken to account for differences that may be critical, e.g., differences in movement capability (such as professional dancer versus a stroke survivor), morphological differences (such as people who use wheelchairs and those who do not). To account for interindividual differences, each data subject could further be treated as a sub-setting, similar to Reference [94].

4.3.5 Contextual Information. Data such as transcript of speech in conversation, musical structure and lyrics of song performed are important contexts for understanding movement behaviour. For example, music performance is often carried out in accordance with a script provided by a musical score, which allows the performers' intentions to be inferred [30]. Other forms of contextual information (e.g., demographic characteristics, personality styles, previous experience, affiliation between interactants, level of cohesion or belonging) can be important predictors to account for. They can also be critical for understanding the extent to which developed models generalize, which is vital for models that, for example, inform clinical practice or are used for automated assessment and decision making. Contextual information can be useful in auditing such models, e.g., to check for bias and discrimination. Existing open movement datasets contain very limited relevant contextual information. While data minimisation, i.e., collection of only the minimum data needed to address a given research question, is good practice, context can play a huge role both in the science and ethics of movement-related research.

How much context needs to be captured? As is the norm, the primary purpose of data collection will usually drive the selection of contextual information to be recorded. However, increasing interest in making data available for secondary use highlights the need for data creators to look toward additional opportunities for use when deciding the context to record. What is even more critical is the need for data creators to provide detailed specification (including the assumptions that they rely on, the context in which the data was captured, and any skew in demographics) of their datasets as standard practice. This is important to ensure that secondary uses respect the limitations of the datasets, although targeting dataset creators for machine learning in the industry [35] contains guidelines and examples that can be useful to the wider research community for creating such documentations.

4.4 Ethics

Although we decided not to include ethics and data protection information in our survey charting, we consider them significant and, more importantly, critical to secondary use of open datasets. Thus, we include a dimension for ethics and data protection in our discussion here as well as in the framework in Table 3. This is because limited ethics and data protection information is a barrier to secondary use of datasets. While the framework for documenting data proposed by Reference [35] (mentioned in Section 4.3.5) is comprehensive and includes ethics-relevant information, there is the need to further emphasise ethical considerations for researchers and across disciplines. This is because sharing of movement data among researchers can be helped or hindered by the amount of discussion of relevant research ethics issues that are considered in the data collection and processing stages prior to release. As ethics approval procedures vary widely across institutions and national contexts, particularly with respect to the use of secondary data, it is critical that for each dataset clear descriptions regarding common ethical concerns are made readily available.

Some of the important questions that such ethics information needs to cover (not at all an exhaustive list) are:

- Was the investigation approved by a relevant ethical approval body (e.g., REC/IRB at a university, or a relevant national body), and if so, is there an approval number that allows the full ethics approval to be consulted?
- Were there any relevant data protection laws involved, and how were these aspects managed (e.g., for those affected by the European Union's General Data Protection Regulation, were privacy notices provided?)? Were any particular codes of research ethics conduct followed

and if so, which versions and when were the data captured? It is important to have this temporal context, since ethical sensitivities change over time and place, and it is important for contextualising prior work when determining the ethics of use at a later time.

- Did all participants give informed consent to the capture, processing, and sharing of their data in the forms offered? Were they alerted to any risks of such sharing (e.g., identity or health conditions being revealed implicitly)? For what specific purposes (e.g., to evaluate computational models for movement during music performances) did they consent to use of their data?
- Was appropriate consideration given to the cultural or personal context and/or meaning of the activities observed for the purpose of data capture? This needs to be documented to ensure that subsequent use of the data maintains integrity with the full context of collection, thereby respecting the participants and their values. This is particularly important when analysing already-secondary data sources such as television programmes or online videos, where the participants themselves may not have explicitly consented to the research.
- Were the participants professionals (e.g., actors, dancers), members of the public, and/or potentially vulnerable (e.g., patients)? What are the age ranges covered in the dataset?
- For data that have been processed to minimise the risk of identification, was this a process of pseudonymisation, de-identification, or full anonymisation? If either of the former, is the potential risk of re-identification described? Were participants made aware of the risk?

While this level of detail may perhaps seem excessive in the context of providing a dataset, it can make a significant difference to the ease with which a potential user can subsequently justify their use of that data to their oversight body. With the variance in oversight criteria on an international scale, the more information that is made available, the easier it is for a researcher to pick out the information required and make their case. Our recommendation, therefore, is that where possible, movement datasets are accompanied by a description of the ethical issues considered before, during, and after data collection (including contextual matters), along with any process-related information such as approval numbers, ethics codes used, and oversight bodies involved. There have been similar discussions on dataset ethics in other contexts (e.g., software repository mining [38]).

4.5 Call for a Unified Human Movement Bank

As our findings of the distributions of open datasets across the years show (see the introduction in Section 3), datasets are increasingly becoming recognised as valuable research contribution, and open data is a culture that is expanding across the research community [64, 84]. This makes it especially timely to create a common repository for human movement datasets across research groups, countries, and disciplines.

Bricks of a general repository already exist across several local initiatives, for instance in dance annotation and choreography (e.g., motionbank.org), in skill acquisition (e.g., Reference [11]), and surveys like ours. However, a systematic and homogeneous format(s) of dataset descriptions across all those datasets is missing. Relevant examples of model international databases to follow exist in various research fields. One example is the Physionet database (physionet.org), since the early 1980s offering to the biomedical community free access to large collections of physiological and clinical data and related open-source software. Another and more recent example is the EBRAINS digital research infrastructure (ebrains.eu) created by the EU-funded Human Brain project, which gathers an extensive range of data and tools for brain-related research. Other burgeoning initiatives within AI communities offering software, data, and practical tools include the UC Irvine Machine Learning Repository (archive.ics.uci.edu/ml), the European AI on Demand

Platform (www.ai4europe.eu), and the HumaneAI network (www.humane-ai.eu) built to promote the development and benchmarking of AI systems.

Following these and other initiatives, we suggest that the *Human Movement Bank* should include:

- an extensive archive of digital recordings, with detailed specifications, of human movement (and para-movement) signals and relevant annotations for use by our research communities,
- a registry of model data collection protocols that can be used for replicating and reproducing datasets, e.g., to increase the data size for an existing dataset, or to reproduce an existing dataset for a different age group or culture,
- a software library of classic and contemporary signal processing and analysis toolboxes as well as machine learning algorithms and pre-trained models for movement data, and
- an ensemble of tutorials and other educational materials.

Although we could not already present the building blocks (e.g., a website that dataset owners can already add their datasets to) for such a repository in the current article, discussions on next steps toward realizing it are already underway. Here, we merely aimed to start a conversation on it within the research community. One challenge that we foresee is in the operationalisation of a common data specification format for all datasets. We hope to learn from similar endeavors in accumulating, e.g., large-scale, open access databases of brain imaging data over the past two decades encouraged by scientific advances associated with similar ventures in the field of genomics [82]. Consistent file organization standards and sufficient quantities of data have proven necessary to pushing forward this initiative [59, 60]. One of the significant outcomes reaped has been advances in methods development (e.g., new computational tools for examining communication within and between brain networks) and studies of individual differences [59], notably in the discovery of behavioral phenotypes by examining relations between patterns of population variability in the brain and performance on a range of tasks (e.g., Reference [39]).

It is important to recognise that concentrating research data in this way may create new ethical issues as a result of that very concentration. A single apparently authoritative and comprehensive data source could create further homogeneity in research by virtue of researchers opting to use it rather than collect their own data. While good in terms of data reuse and lowering the collective burden of research on participants, there is the risk that biases or omissions in the available data are perpetuated through reuse rather than corrected through new acquisition. It is therefore critical that specifications for individual datasets on the proposed repository include characterisation of their limitations. Reviews of the kind provided in this article will further be important to give an overview of bias and omission in the whole collection. This would help to offset the increased risk of perpetuating the use of less inclusive data by using the power of the combined data to identify opportunities for more equitable and inclusive data acquisition in future.

5 LIMITATIONS

We have presented our 704-item catalogue of open human movement datasets. For each dataset, we provide information as a starting point for researchers to find datasets relevant to their research interests. We additionally contribute analyses of the datasets with respect to pertinent attributes that cover motivation, dataset contents, and subject population groups. We also reviewed the datasets along the themes of diversity, ecological validity, and data recorded, highlighting a 12-dimension framework for dataset creators to use in planning and building data corpora. Nevertheless, we note a few limitations of our study.

Although our aim was to provide a survey of human movement datasets as comprehensive as possible, we are aware that eligible datasets are likely missing from our final list. This is primarily due to limitations of digital archiving of research articles. For example, metadata such as article title and keywords did not always contain information that highlighted that the corresponding article was the primary reference for a dataset. This is possibly because datasets have not always been recognised as significant research contributions [72]. Meanwhile, the universality of the term "data" or "dataset" and data itself meant that searching by main text was not an efficient approach, as it resulted in unwieldly output. This can be evidenced in the outcome of our search within the main text (and not just the titles and keywords) in the *SpringerLink* digital library. Even with fine-tuning (and automatic capping done by the library's search system), our search resulted in just about 40,000 relevant articles eligible for the first-level screening. It took three full months to complete the first-level screening for that set of results. Given the time criticality of a scoping review, it was not feasible to use the same approach for other article databases.

Another relevant challenge that we faced in indexing datasets concerns the level of information provided in research articles returned by our search. While a large number provided descriptive information that we were able to use for second-level screening and charting, several of them either did not include sufficient information to discern whether or not they met the primary eligibility criterion (i.e., that they referred to human movement data), had little information that we could chart, or did not make clear if the dataset was open for reuse by the research community. We did notice a few good models of dataset records that clearly described the data. However, it is clear that such practices need to become the norm rather than exceptions. As highlighted in Section 4.3.5, the importance of this has recently and at several occasions been flagged in the machine learning community. Detailed and structured metadata for datasets will enable more critical analyses of individual and collective research findings and outcomes that is important to scientific knowledge and technology development. It will, of course, also be vital for finding open datasets and reusing them. In their paper on this topic, Reference [35] proposed what they referred to as "datasheets" for datasets. Such technical specifications would include motivation information such as funding source and creation purpose. It will also include extensive details about what the dataset contains, how it was collected, and what it can be used for. Importantly, as highlighted in Section 4.4, information about ethics would additionally be included. A significant characteristic of the proposed datasheet is its clear structure.

Last, we note that our eligibility screening and charting was done by one person and without independent testing for reliability. Nevertheless, we will be making our final survey outcome (the list of 704 human movement datasets) publicly available for researchers to contribute to.

6 CONCLUSION

We curated a list of 704 human movement datasets available for secondary use within the research community (see supplementary material) and used it to characterise the current coverage that such datasets offer to researchers. This movement dataset landscape has been described in terms of 10 basic variables: name and reference, creation purpose, data type, annotations, source, population groups, ordinal size of people captured simultaneously, URL, motion capture sensor, and funders. Like all reviews, our results will eventually become outdated, but we expect that they will for a longer time remain a (historical) reference for the community.

Further, we contribute a framework that may help researchers creating new movement datasets in considering important factors that: (1) drive and affect movement, (2) may lead to a more inclusive and ecologically valid understanding of movement, and (3) can support the sharing of data, which is an expensive and important resource. These are timeless themes influential in the advance of scientific knowledge and technological development related to human movement. We do not consider this framework complete but rather a starting point for collaborating across disciplines.

REFERENCES

[1] Md Atiqur Rahman Ahad. 2014. Datasets for action, gesture and activity analysis. In *2nd International Conference on Intelligent Systems and Image Processing*.

[2] Md Atiqur Rahman Ahad, J. Tan, H. Kim, and S. Ishikawa. 2011. Action dataset—A survey. In *SICE Annual Conference*. IEEE, 1650–1655.

[3] AniAgeProjectdataset. (????). Traditional Dances Download. Retrieved from https://www.euh2020aniage.org/testthaidancedownload.

[4] Edouard Auvinet, Franck Multon, Alain Saint-Arnaud, Jacqueline Rousseau, and Jean Meunier. 2010. Fall detection with multiple cameras: An occlusion-resistant method based on 3-D silhouette vertical distribution. *IEEE Trans. Inf. Technol. Biomed.* 15, 2 (2010), 290–300.

[5] Giulia Barbareschi and D. Morgado Ramirez. 2020. Supporting the voice of people with disabilities in Kenya, Uganda and Jordan. In *Rethinking Giving Voice Workshop*, Vol. 2020. Association for Computing Machinery.

[6] Benoît G. Bardy, Carmela Calabrese, Pietro De Lellis, Stella Bourgeaud, Clémentine Colomer, Simon Pla, and Mario di Bernardo. 2020. Moving in unison after perceptual interruption. *Sci. Rep.* 10, 1 (2020), 1–13.

[7] Lisa Feldman Barrett. 2017. The theory of constructed emotion: An active inference account of interoception and categorization. *Soc. Cog. Affect. Neurosci.* 12, 1 (2017), 1–23.

[8] Lisa Feldman Barrett and W. Kyle Simmons. 2015. Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16, 7 (2015), 419–429.

[9] David R. Bassett, John Pucher, Ralph Buehler, Dixie L. Thompson, and Scott E. Crouter. 2008. Walking, cycling, and obesity rates in Europe, North America, and Australia. *J. Phys. Activ. Health* 5, 6 (2008), 795–814.

[10] Roy F. Baumeister and Mark R. Leary. 1995. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 3 (1995), 497.

[11] Massimo Bergamasco, Benoit Bardy, and Daniel Gopher. 2012. *Skill Training in Multimodal Virtual Environments*. CRC Press.

[12] Richa Bharadwaj, Srijittra Swaisaenyakorn, Clive G. Parini, John C. Batchelor, and Akram Alomainy. 2017. Impulse radio ultra-wideband communications for localization and tracking of human body and limbs movement for healthcare applications. *IEEE Trans. Ant. Propag.* 65, 12 (2017), 7298–7309.

[13] Roberta Bianco, Giacomo Novembre, Peter E. Keller, Arno Villringer, and Daniela Sammler. 2018. Musical genre-dependent behavioural and EEG signatures of action planning. A comparison between classical and jazz pianists. *Neuroimage* 169 (2018), 383–394.

[14] M. M. N. Bieńkiewicz, Andrii Smykovskyi, Temitayo Olugbade, Stefan Janaqi, Antonio Camurri, Nadia Bianchi-Berthouze, Mårten Björkman, and Benoît G. Bardy. 2021. Bridging the gap between emotion and joint action. *Neurosci. Biobehav. Rev.* (2021).

[15] Ziyun Cai, Jungong Han, Li Liu, and Ling Shao. 2017. RGB-D datasets using Microsoft Kinect or similar sensors: A survey. *Multim. Tools Applic.* 76, 3 (2017), 4313–4355.

[16] Qiong Cao, Li Shen, Weidi Xie, Omkar M. Parkhi, and Andrew Zisserman. 2018. VGGFace2: A dataset for recognising faces across pose and age. In *13th IEEE International Conference on Automatic Face & Gesture Recognition (FG'18)*. IEEE, 67–74.

[17] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. 2019. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Trans. Patt. Anal. Mach. Intell.* (2019).

[18] Patrick Carrington, Kevin Chang, Helena Mentis, and Amy Hurst. 2015. "But, I don't take steps." Examining the inaccessibility of fitness trackers for wheelchair athletes. In *17th International ACM Conference on Computers & Accessibility*. 193–201.

[19] Andrea Cavallo, Nathan C. Foster, Karthikeyan Kalyanasundaram Balasubramanian, Andrea Merello, Giorgio Zini, Marco Crepaldi, and Cristina Becchio. 2021. A low-cost stand-alone platform for measuring motor behaviour across developmental applications. *iScience* (2021), 102742.

[20] Jose M. Chaquet, Enrique J. Carmona, and Antonio Fernández-Caballero. 2013. A survey of video datasets for human action and activity recognition. *Comput. Vis. Image Underst.* 117, 6 (2013), 633–659. DOI: https://doi.org/10.1016/j.cviu.2013.01.013

[21] Youngjun Cho, Nadia Bianchi-Berthouze, and Simon J. Julier. 2017. DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In *7th International Conference on Affective Computing and Intelligent Interaction*. IEEE, 456–463.

[22] Martin Clayton, Kelly Jakubowski, Tuomas Eerola, Peter E. Keller, Antonio Camurri, Gualtiero Volpe, and Paolo Alborno. 2020. Interpersonal entrainment in music performance: Theory, method, and model. *Mus. Percept.: Interdisc. J.* 38, 2 (2020), 136–194.

[23] Martin Clayton, Laura Leante, and Simone Tarsitani. 2018. IEMP North Indian Raga. DOI: https://doi.org/10.17605/OSF.IO/KS325

[24] A. Cook, B. Mandal, D. Berry, and M. Johnson. 2019. Towards automatic screening of typical and atypical behaviors in children with autism. In *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. DOI: https://doi.org/10.1109/DSAA.2019.00065

[25] Emiro De-La-Hoz-Franco, Paola Ariza-Colpas, Javier Medina Quero, and Macarena Espinilla. 2018. Sensor-based datasets for human activity recognition–a systematic review of literature. *IEEE Access* 6 (2018), 59192–59210.

[26] Koldo De Miguel, Alberto Brunete, Miguel Hernando, and Ernesto Gambao. 2017. Home camera-based fall detection system for the elderly. *Sensors* 17, 12 (2017), 2864.

[27] Ezequiel A. Di Paolo, Elena Clare Cuffari, and Hanne De Jaegher. 2018. *Linguistic Bodies: The Continuity between Life and Language*. The MIT Press.

[28] Philippe Dreuw, Jens Forster, and Hermann Ney. 2010. Tracking benchmark databases for video-based sign language recognition. In *European Conference on Computer Vision*. Springer, 286–297.

[29] Bernd Dudzik, Michel Pierre Jansen, Franziska Burger, Frank Kaptein, Joost Broekens, Dirk K. J. Heylen, Hayley Hung, Mark A. Neerincx, and Khiet P. Truong. 2019. Context in human emotion perception for automatic affect detection: A survey of audiovisual databases. DOI: https://doi.org/10.1109/ACII.2019.8925446

[30] Alessandro D'Ausilio, Giacomo Novembre, Luciano Fadiga, and Peter E. Keller. 2015. What can music tell us about social interaction? *Trends. Cog. Sci.* 19, 3 (2015), 111–114.

[31] Michael Edwards, Jingjing Deng, and Xianghua Xie. 2016. From pose to activity: Surveying datasets and introducing CONVERSE. *Comput. Vis. Image Underst.* 144 (2016), 73–105.

[32] Robert Wall Emerson. 2015. Convenience sampling, random sampling, and snowball sampling: How does sampling affect the validity of research? *J.Vis. Impair. Blind.* 109, 2 (2015), 164–168.

[33] Michael Firman. 2016. RGBD datasets: Past, present and future. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 19–31.

[34] Vittorio Gallese. 2007. Before and below "theory of mind": Embodied simulation and the neural correlates of social cognition. *Philos. Trans. Roy. Societ. B: Biol. Sci.* 362, 1480 (2007), 659–669. DOI: https://doi.org/10.1098/rstb.2006.2002

[35] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé, and Kate Crawford. 2018. Datasheets for datasets. arXiv (2018).

[36] Apostolos P. Georgopoulos. 2000. Neural aspects of cognitive motor control. *Curr. Opin. Neurobiol.* 10, 2 (2000), 238–241.

[37] Kathrin Gerling and Katta Spiel. 2021. A critical examination of virtual reality technology in the context of the minority body. In *CHI Conference on Human Factors in Computing Systems*. 1–14.

[38] Nicolas E. Gold and Jens Krinke. 2022. Ethics in the mining of software repositories. *Empir. Softw. Eng.* 27, 1 (2022), 1–49.

[39] Weikang Gong, Christian F. Beckmann, and Stephen M. Smith. 2021. Phenotype discovery from population brain imaging. *Med. Image Anal.* 71 (2021), 102050.

[40] Chunhui Gu, Chen Sun, David A. Ross, Carl Vondrick, Caroline Pantofaru, Yeqing Li, Sudheendra Vijayanarasimhan, George Toderici, Susanna Ricco, Rahul Sukthankar, et al. 2018. AVA: A video dataset of spatio-temporally localized atomic visual actions. In *IEEE Conference on Computer Vision and Pattern Recognition*. 6047–6056.

[41] Hatice Gunes and Massimo Piccardi. 2006. Creating and annotating affect databases from face and body display: A contemporary survey. In *IEEE International Conference on Systems, Man and Cybernetics*, Vol. 3. 2426–2433. DOI: https://doi.org/10.1109/ICSMC.2006.385227

[42] Hatice Gunes, Caifeng Shan, Shizhi Chen, and YingLi Tian. 2015. Bodily expression for automatic affect recognition. *Emot. Recog.: Patt. Anal. Appr.* (2015), 343–377.

[43] Tal Hassner. 2013. A critical review of action recognition benchmarks. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. 245–250. DOI: https://doi.org/10.1109/CVPRW.2013.43

[44] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.

[45] Joseph Henrich, Steven J. Heine, and Ara Norenzayan. 2010. The weirdest people in the world? *Behav. Brain Sci.* 33, 2-3 (2010), 61–83.

[46] Ryan M. Hulteen, Jordan J. Smith, Philip J. Morgan, Lisa M. Barnett, Pedro C. Hallal, Kim Colyvas, and David R. Lubans. 2017. Global participation in sport and leisure-time physical activities: A systematic review and meta-analysis. *Prevent. Med.* 95 (2017), 14–25.

[47] Marco Iacoboni, Istvan Molnar-Szakacs, Vittorio Gallese, Giovanni Buccino, and John C. Mazziotta. 2005. Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3, 3 (2005), 0529–0535. DOI: https://doi.org/10.1371/journal.pbio.0030079

[48] INVOLVE. 2015. *Public Involvement in Research: Values and Principles Framework*. Technical Report.

[49] Luis Jure, Martín Rocamora, Simone Tarsitani, and Martin Clayton. 2018. IEMP Uruguayan Candombe. DOI: https://doi.org/10.17605/OSF.IO/WFX7K

[50] Michelle Karg, Ali-Akbar Samadani, Rob Gorbet, Kolja Kühnlenz, Jesse Hoey, and Dana Kulić. 2013. Body movements for affective expression: A survey of automatic recognition and generation. *IEEE Trans. Affect. Comput.* 4, 4 (2013), 341–359.

[51] Peter E. Keller, Giacomo Novembre, and Michael J. Hove. 2014. Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philos. Trans. Roy. Societ. B: Biol. Sci.* 369, 1658 (2014), 20130394.

[52] Sohaib Khan and Mubarak Shah. 2000. Tracking people in presence of occlusion. In *Asian Conference on Computer Vision*, Vol. 5. Citeseer.

[53] Kyungwan Kim and Otmar Bock. 2019. Ecological validity of manual grasping movements in an everyday-like grocery shopping task. *Experim. Brain Res.* 237, 5 (2019), 1169–1177.

[54] Rob Kitchin. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences.* London. DOI: https://doi.org/10.4135/9781473909472

[55] Andrea Kleinsmith and Nadia Bianchi-Berthouze. 2013. Affective body expression perception and recognition: A survey. *IEEE Trans. Affect. Comput.* 4, 1 (2013), 15–33.

[56] NIHR Research Design Service London. 2018. *Patient and Public Involvement in Health and Social Care Research: A Handbook for Researchers.* Technical Report. Retrieved from https://www.rds-london.nihr.ac.uk/wpcms/wp-content/uploads/2018/10/RDS_PPI-Handbook_2018_WEB_VERSION.pdf.

[57] Yu Luo, Jianbo Ye, Reginald B. Adams, Jia Li, Michelle G. Newman, and James Z. Wang. 2020. ARBEE: Towards automated recognition of bodily expression of emotion in the wild. *Int. J. Comput. Vis.* 128, 1 (2020), 1–25.

[58] Vincenzo Lussu, Radoslaw Niewiadomski, Gualtiero Volpe, and Antonio Camurri. 2020. The role of respiration audio in multimodal analysis of movement qualities. *J. Multim. User Interf.* 14, 1 (2020), 1–15.

[59] Christopher R. Madan. 2021. Scan once, analyse many: Using large open-access neuroimaging datasets to understand the brain. *Neuroinformatics* (2021), 1–29.

[60] Scott Marek, Brenden Tervo-Clemmens, Finnegan J. Calabro, David F. Montez, Benjamin P. Kay, Alexander S. Hatoum, Meghan Rose Donohue, William Foran, Ryland L. Miller, Timothy J. Hendrickson, et al. 2022. Reproducible brain-wide association studies require thousands of individuals. *Nature* 603, 7902 (2022), 654–660.

[61] Michał Meina, Andrzej Janusz, Krzysztof Rykaczewski, Dominik Ślęzak, Bartosz Celmer, and Adam Krasuski. 2015. Tagging firefighter activities at the emergency scene: Summary of AAIA'15 data mining competition at knowledge Pit. In *Federated Conference on Computer Science and Information Systems (FedCSIS).* IEEE, 367–373.

[62] Dirk Moelants, Michiel Demey, Maarten Grachten, Chia-Fen Wu, and Marc Leman. 2012. The influence of an audience on performers: A comparison between rehearsal and concert using audio, video and movement data. *J. New Mus. Res.* 41, 1 (2012), 67–78.

[63] Zachary Munn, Micah D. J. Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med. Res. Methodol.* 18, 1 (2018), 1–7. DOI: https://doi.org/10.1186/s12874-018-0611-x

[64] Peter Murray-Rust. 2008. Open data in science. *Serials Rev.* 34, 1 (2008), 52–64. DOI: https://doi.org/10.1016/j.serrev.2008.01.001

[65] Fatemeh Noroozi, Dorota Kaminska, Ciprian Corneanu, Tomasz Sapinski, Sergio Escalera, and Gholamreza Anbarjafari. 2018. Survey on emotional body gesture recognition. *IEEE Trans. Affect. Comput.* (2018).

[66] João Ferreira Nunes, Pedro Miguel Moreira, and João Manuel R. S. Tavares. 2019. Benchmark RGB-D gait datasets: A systematic review. In *ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing*. Springer, 366–372.

[67] Temitayo A. Olugbade, Aneesha Singh, Nadia Bianchi-Berthouze, Nicolai Marquardt, Min S. H. Aung, and Amanda C. De C Williams. 2019. How can affect be detected and represented in technological support for physical rehabilitation? *ACM Trans. Comput.-Hum. Interact.* 26, 1 (2019), 1–29.

[68] Rainer Polak, Simone Tarsitani, and Martin Clayton. 2018. IEMP Malian Jembe. DOI: https://doi.org/10.17605/OSF.IO/M652X

[69] Sreenivasan Ramasamy Ramamurthy and Nirmalya Roy. 2018. Recent trends in machine learning for human activity recognition–A survey. *Data Mining Knowl. Discov.* 8, 4 (2018), e1254.

[70] Simon Ruffieux, Denis Lalanne, Elena Mugellini, and Omar Abou Khaled. 2014. A survey of datasets for human gesture recognition. In *International Conference on Human-computer Interaction.* Springer, 337–348.

[71] Nithya Sambasivan, Erin Arnesen, Ben Hutchinson, Tulsee Doshi, and Vinodkumar Prabhakaran. 2021. Re-imagining algorithmic fairness in India and beyond. In *ACM Conference on Fairness, Accountability, and Transparency.* 315–328.

[72] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M. Aroyo. 2021. "Everyone wants to do the model work, not the data work": Data cascades in high-stakes AI. In *CHI Conference on Human Factors in Computing Systems.* 1–15.

[73] Leonhard Schilbach, Bert Timmermans, Vasudevi Reddy, Alan Costall, Gary Bente, Tobias Schlicht, and Kai Vogeley. 2013. Toward a second-person neuroscience 1. *Behav. Brain Sci.* 36, 4 (2013), 393–414.

[74] Maxine Sheets-Johnstone. 2011. *The Primacy of Movement. (Advances in Consciousness Research, Vol. 82.)* John Benjamins Publishing Company, Amsterdam. DOI: https://doi.org/10.1075/aicr.82

[75] Roshan Singh, Ankur Sonawane, and Rajeev Srivastava. 2020. Recent evolution of modern datasets for human activity recognition: A deep survey. *Multim. Syst.* 26, 2 (2020), 83–106.

[76] Tej Singh and Dinesh Kumar Vishwakarma. 2019. Video benchmarks of human action datasets: A review. *Artif. Intell. Rev.* 52, 2 (2019), 1107–1154.

[77] Katta Spiel. 2021. The bodies of TEI—Investigating norms and assumptions in the design of embodied interaction. In *International Conference on Tangible, Embedded, and Embodied Interaction.* 1–19.

[78] Benjamin Stephens-Fripp, Fazel Naghdy, David Stirling, and Golshah Naghdy. 2017. Automatic affect perception based on body gait and posture: A survey. *Int. J. Soc. Robot.* 9, 5 (2017), 617–641.

[79] Lena H. Ting, Hillel J. Chiel, Randy D. Trumbower, Jessica L. Allen, J. Lucas McKay, Madeleine E. Hackney, and Trisha M. Kesar. 2015. Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron* 86, 1 (2015), 38–54.

[80] Andrea C. Tricco, Erin Lillie, Wasifa Zarin, Kelly K. O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah D. J. Peters, Tanya Horsley, Laura Weeks, Susanne Hempel, Elie A. Akl, Christine Chang, Jessie McGowan, Lesley Stewart, Lisa Hartling, Adrian Aldcroft, Michael G. Wilson, Chantelle Garritty, Simon Lewin, Christina M. Godfrey, Marilyn T. MacDonald, Etienne V. Langlois, Karla Soares-Weiser, Jo Moriarty, Tammy Clifford, Özge Tunçalp, and Sharon E. Straus. 2018. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann. Internal Med.* 169, 7 (2018), 467–473. DOI: https://doi.org/10.7326/M18-0850

[81] Melissa A. Valerio, Natalia Rodriguez, Paula Winkler, Jaime Lopez, Meagen Dennison, Yuanyuan Liang, and Barbara J. Turner. 2016. Comparing two sampling methods to engage hard-to-reach communities in research priority setting. *BMC Med. Res. Methodol.* 16, 1 (2016), 1–11.

[82] John D. Van Horn and Michael S. Gazzaniga. 2002. Databasing fMRI studies—Towards a "discovery science" of brain function. *Nat. Rev. Neurosci.* 3, 4 (2002), 314–318.

[83] Somogy Varga and Detlef H. Heck. 2017. Rhythms of the body, rhythms of the brain: Respiration, neural oscillations, and embodied cognition. *Conscious. Cogn.* 56 (2017), 77–90.

[84] Todd J. Vision. 2010. Open data and the social contract of scientific publishing. *BioScience* 60, 5 (2010), 330–331. DOI: https://doi.org/10.1525/bio.2010.60.5.2

[85] Athanasios Voulodimos, Dimitrios Kosmopoulos, Georgios Vasileiou, Emmanuel Sardis, Vasileios Anagnostopoulos, Constantinos Lalos, Anastasios Doulamis, and Theodora Varvarigou. 2012. A threefold dataset for activity and workflow recognition in complex industrial environments. *IEEE MultiM.* 19, 03 (2012), 42–52.

[86] J. C. K. Wells, T. J. Cole, D. Bruner, and P. Treleaven. 2008. Body shape in American and British adults: Between-country and inter-ethnic comparisons. *Int. J. Obes.* 32, 1 (2008), 152–159.

[87] Meredith Whittaker, Meryl Alper, Cynthia L. Bennett, Sara Hendren, Liz Kaziunas, Mara Mills, Meredith Ringel Morris, Joy Rankin, Emily Rogers, Marcel Salas, et al. 2019. *Disability, Bias, and AI.* Technical Report.

[88] Rua M. Williams and Juan E. Gilbert. 2019. Cyborg perspectives on computing research reform. In *CHI Conference on Human Factors in Computing Systems.* 1–11.

[89] Shihao Xu, Jing Fang, Xiping Hu, Edith Ngai, Yi Guo, Victor Leung, Jun Cheng, and Bin Hu. 2020. Emotion recognition from gait analyses: Current research and future directions. *arXiv preprint arXiv:2003.11461* (2020).

[90] Jin H. Yan, Jerry R. Thomas, George E. Stelmach, and Katherine T. Thomas. 2000. Developmental features of rapid aiming arm movements across the lifespan. *J. Motor Behav.* 32, 2 (2000), 121–140.

[91] Haris Zacharatos, Christos Gatzoulis, and Yiorgos L. Chrysanthou. 2014. Automatic emotion recognition based on body movement analysis: A survey. *IEEE Comput. Graph. Applic.* 34, 6 (2014), 35–45.

[92] Jing Zhang, Wanqing Li, Philip O. Ogunbona, Pichao Wang, and Chang Tang. 2016. RGB-D-based action recognition datasets: A survey. *Patt. Recog.* 60 (2016), 86–105.

[93] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. 2020. A comprehensive survey on transfer learning. *Proc. IEEE* 109, 1 (2020), 43–76.

[94] Andrea Zunino, Jacopo Cavazza, Riccardo Volpi, Pietro Morerio, Andrea Cavallo, Cristina Becchio, and Vittorio Murino. 2020. Predicting intentions from motion: The subject-adversarial adaptation approach. *Int. J. Comput. Vis.* 128, 1 (2020), 220–239.

Received 4 January 2022; revised 27 April 2022; accepted 2 May 2022